The Bone & Joint Journal

Hosman AJF, Barbagallo G, the SCI-POEM Study Group, van Middendorp JJ. Neurological recovery after early versus delayed surgical decompression for acute traumatic spinal cord injury. *Bone Joint J*. 2023;105-B(4):400-411.

https://doi.org/10.1302/0301-620X.105B4.BJJ-2022-0947.R2

11 May 2023

Authors' reply:

Sir,

We would like to thank Dr ter Wengel et al for their interest in our paper[1] and we welcome the opportunity to clarify the concerns they raise. Each of these clarifications had already been addressed throughout the writing process of the manuscript, yet we are grateful to the senior author of the letter for the opportunity to share these in the public domain.

First, we have to correct a misinterpretation of our study: the primary aim of our study was not to create similar groups of patients treated in different timeframes. Instead, we primarily aimed to generate a dataset that would enable us to reject the null hypothesis, i.e. early (≤ 12 hours) surgical spinal decompression does not result in a better neurological recovery 12 months after injury compared with late (> 12 hours and < 14 days) surgical spinal decompression in patients with acute traumatic spinal cord injury (tSCI). As extensively covered throughout our paper, we were very mindful of potential baseline imbalances in our non-randomized study from the inception through to the analytical phase of this decade-long project. Hence, to the best of our knowledge, we are the first in this area of research to account for the expected heterogeneity in this population by the use of sophisticated analytical techniques, per protocol.

Second, Dr ter Wengel et al seem to assign more value to unadjusted analyses than adjusted analyses. For instance, "patients with tetraplegia in the early cohort had significantly lower LEMS at baseline (10.5 versus 25.3) and experienced 16.7 LEMS recovery versus 14.0 in the late cohort" followed by an explanation on the functional implications of such differences in recovery. They fail to recognize, however, that the ceiling effect, which we clearly explained in the discussion section, possibly limited the potential for recovery beyond 14.0 LEMS points in the late cohort. This is a critically important discussion point as these authors and many others – including the Surgical Timing in Acute Spinal Cord Injury Study (STASCIS) investigators[2] – have found comfort in showing (post hoc) unadjusted analyses favouring early surgical decompression, without adjusting for complex, neurological intricacies seen in patients with tSCI. This tendency may well explain the deeply rooted publication bias within this field on which we were the first to report a decade ago.[3] The very same meta-analysis also covered and contests Dr ter Wengel et al's statement on the era of 'equipoise'.[3] In order to arrive at balanced, hypothesis-generating conclusions on secondary outcomes, including outcomes in various tSCI subgroups, exploratory analyses will be covered in a secondary paper as referred to in our discussion.

Third, Dr ter Wengel et al indicated that we "did not acknowledge that the recovery of LEMS is highly variable depending on the initial severity of neurological injury". In our methods section, we clearly outlined which variables were considered for generating the propensity score: these include multiple measures of injury severity at baseline. They then commented that too many variables were considered for the propensity score model. For developing the propensity score model, we have followed recommendations from published literature. We have selected the final model based on balance diagnostics, i.e. we have used the model that achieved the best possible balance given the observed baseline imbalance.[4] Furthermore, we have included variables that have been shown to be strong predictors of the outcome, i.e. their inclusion has been shown to lead to increased precision of the treatment effect.[5] Concerning the observed residual imbalance in the baseline AIS grade, we would like to highlight that this has been addressed by including this factor as one of the variables in the outcome regression model.

We agree with Dr ter Wengel et al that in general there may be a bias-variance trade-off when defining a propensity score model. However, we think we can allay their concerns that the inclusion of a large number of variables in our propensity score model might have negatively affected, i.e. increased, the variance of the estimated treatment effect. The variance of the estimated treatment effect is directly reflected in the estimate's confidence interval. Compared with the results from the unadjusted analysis (confidence interval (-0.3 to 8.8)), our propensity score-adjusted analysis yielded a narrower confidence interval ranging from -1.5 to 5.9. This demonstrates that the predictors included in our propensity score model have improved the precision over the unadjusted analysis. Furthermore, our study was powered to detect a difference in LEMS of six points or greater. The confidence interval of the adjusted analysis does not include the mentioned value and hence we can exclude a treatment effect of that size or larger which confirms that the selected propensity score model has not negatively affected study power.

Last, Dr ter Wengel et al are correct that pre- and postoperative imaging will provide important additional information to our analyses, which we indeed did outline as a (fourth) study limitation in the discussion section of the paper. The attrition rate and related missing data were also acknowledged as a second limitation. Attrition, along with the relative impact of injury level and severity, is a well-established challenge in clinical SCI research.[6] In our sample size estimation, we assumed an overall dropout or incomplete data rate of 30%. Knowing that minimizing attrition was a key challenge throughout the conduct of the study and considering that both trauma and rehabilitation centres were involved in individual patient management, we were genuinely pleased with our relatively high rate of follow-up.

On behalf of the SCI-POEM Study Group,
A. J. F. Hosman, MD, PhD,
Department of Orthopaedic Surgery, Radboud University Medical Center,
Nijmegen, The Netherlands.
G. Barbagallo, MD,
U.O. di Clinica Neurochirurgia Azienda Ospedaliero – Universitaria Policlinico,

Catania, Italy.
J. J. van Middendorp, MD, PhD,
Department of Orthopaedic Surgery, Radboud University Medical Center,
Nijmegen, The Netherlands.

1. **Hosman AJF, Barbagallo G, the SCI-POEM Study Group, van Middendorp JJ.** Neurological recovery after early versus delayed surgical decompression for acute traumatic spinal cord injury. *Bone Joint J*. 2023;105-B(4):400-411.
2. **van Middendorp JJ.** Letter to the editor regarding: "Early versus delayed decompression for traumatic cervical spinal cord injury: results of the Surgical Timing in Acute Spinal Cord Injury Study (STASCIS)". *Spine J*. 2012;12(6):540; author reply 541-2. doi: 10.1016/j.spinee.2012.06.007. PMID: 22857647.
3. **van Middendorp JJ, Hosman AJF, Doi SAR.** The effects of the timing of spinal surgery after traumatic spinal cord injury: a systematic review and meta-analysis. *J Neurotrauma*. 2013;30(21):1781-1794.
4. **D'Agostino RB Jr.** Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Stat Med*. 1998;17(19):2265-2281.
5. **Brookhart MA, Schneeweiss S, Rothman KJ, Glynn RJ, Avorn J, Stürmer T.** Variable selection for propensity score models. *Am J Epidemiol*. 2006;163(12):1149-1156.
6. **Kim H, Cutter GR, George B, Chen Y.** Understanding and Preventing Loss to Follow-up: Experiences From the Spinal Cord Injury Model Systems. *Top Spinal Cord Inj Rehabil*. 2018;24(2):97-109.